

引文格式:

梁忻健, 姚迪, 文永明, 等. 基于大模型的时态数据分析算法综述 [J]. 集成技术, 2025, 14(5): 1-19.

Liang XJ, Yao D, Wen YM, et al. A survey of temporal data analysis algorithms based on large models [J]. Journal of Integration Technology, 2025, 14(5): 1-19.

基于大模型的时态数据分析算法综述

梁忻健¹ 姚迪^{1*} 文永明² 吕勇梁² 毕经平¹

¹(中国科学院计算技术研究所 北京 100190)

²(北京控制与电子技术研究所 北京 100045)

摘要 时态数据分析通过捕捉时态数据的变化趋势, 揭示数据中潜在的规律和模式, 从而帮助做出精准的预测和决策。近年来, 受大模型强泛化能力的启发, 许多工作将原本用于自然语言处理的大模型拓展至时态数据分析领域, 赋予了时态数据分析模型零样本、多模态推理的能力。但现有工作对基于大模型的时态数据分析算法的系统性分类与讨论较少。因此, 本文首先将现有工作整理成两类: 基于大语言模型的时态数据分析模型和时态基座模型, 并分别按照在大模型工作流程中的优化方向进一步进行分类, 阐述了将大模型应用至时态分析算法的各类方法, 并进行优缺点评估与适用场景分析, 为将大模型嵌入时态分析算法提供了参考; 其次, 总结出基于大模型的时态数据分析模型在零样本推理任务下的性能较强; 再次, 与时态基座模型相比, 由于基于大语言模型的时态分析模型能利用预训练知识, 因此表现出较好的泛化能力, 且计算成本较低, 但分析性能不如时态基座模型; 最后, 本文强调了未来的工作挑战和可能发展方向。

关键词 大模型; 时态数据分析; 基座模型; 模态对齐

中图分类号 TP391.4 文献标志码 A doi: 10.12146/j.issn.2095-3135.20250427001

CSTR: 32239.14.j.issn.2095-3135.20250427001

A Survey of Temporal Data Analysis Algorithms Based on Large Models

LIANG Xinjian¹ YAO Di^{1*} WEN Yongming² LV Yongliang² BI Jingping¹

¹(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)

²(Beijing Institute of Control and Electronic Technology, Beijing 100045, China)

*Corresponding Author: yaodi@ict.ac.cn

Abstract Temporal data analysis helps to make accurate predictions and decisions by capturing the changing trends of temporal data and revealing the underlying rules and patterns. In recent years, inspired by the strong generalization ability of large models, many works have extended the large models originally designed for natural language processing to the field of temporal data analysis, which empowers the temporal

收稿日期: 2025-04-27 修回日期: 2025-07-04

基金项目: 国家自然科学基金项目 (62472405)

作者简介: 梁忻健, 硕士, 研究方向为时空数据挖掘; 姚迪 (通讯作者), 副研究员, 研究方向为时空数据挖掘, E-mail: yaodi@ict.ac.cn; 文永明, 博士, 研究方向为智能决策控制; 吕勇梁, 工程师, 研究方向为智能决策; 毕经平, 研究员, 研究方向为时空数据挖掘。

data analysis models with zero-shot and multimodal inferences. However, there are fewer systematic classifications and discussions on temporal data analysis algorithms based on large models in the existing works. This article first classifies existing works into 2 categories: temporal data analysis models based on large language models and temporal foundation models, further categorizes them respectively according to the directions of targeted optimization in the large model pipeline, describes the various types of methods for applying the large models to the temporal data analysis algorithms, evaluates the strengths and weaknesses and analyzes the applicability scenarios of each type of method, to provide methodological references for applying the large model to the temporal data analysis model. Then it is concluded that the temporal data analysis models based on large models perform strongly under the zero-shot inference task. Because the temporal data analysis models based on large language models can utilize the pre-training knowledge, they exhibit better generalization ability with lower computational cost compared to the temporal foundation model. However, the analysis performance is not as good as that of the temporal foundation model. Finally, the remaining challenges and potential future research directions are highlighted.

Keywords large model; temporal data analysis; foundation model; modality alignment

Funding This work is supported by National Natural Science Foundation of China (62472405)

1 引 言

时态数据指具有时间维度的数据，主要分为时序数据和时空数据。时态数据在现实生活中的应用极其广泛，包括经济^[1]、医疗^[2]、气象^[3]、交通^[4]、能源^[5]和家居^[6]等领域。时态数据分析通过捕捉时态数据的变化趋势，揭示数据中潜在的规律和模式，帮助相关人员做出精准的预测和决策，在各应用领域中都发挥着重要作用。

近年来，得益于深度学习、自然语言处理 (natural language processing, NLP) 和大规模数据的快速进步，大模型在多个方面取得了显著突破。大模型最初主要用于自然语言处理任务，即大语言模型 (large language model, LLM)，如文本生成、情感分析和机器翻译，但随着模型架构和训练方法的进步，大模型的能力得到了显著扩展，展现出强大的泛化能力，不仅能处理文本数据，还在跨领域任务中取得了令人瞩目的成果。

受 LLM 强大的零样本泛化能力启发，通过将大模型与时态数据分析结合，可利用大模型强大的泛化能力解决时态数据分析问题^[7]。

时态数据中的时序数据与时空数据都具有时间特性，均存在隐藏的时间模式信息。因此，结合两种数据进行研究有助于读者更全面地了解基于大模型的时态数据分析。时态数据分析的发展历程可分为 4 个阶段：统计算法、传统机器学习算法、深度学习算法和基于大模型的算法。

在早期，时态数据分析主要依赖于经典的统计模型，最具代表性的模型是自回归积分滑动平均模型。这些模型基于线性假设，对时态数据进行建模，通常用于预测、季节性分析和趋势检测等任务。虽然这些方法相对简单、解释性强，但主要依赖对数据的线性假设，有一定局限性，难以捕获复杂的非线性特征。

随着数据量的增大和计算能力的提升，机器学习方法逐渐被引入时态数据分析^[8]。与传统的

统计模型相比, 机器学习模型 (如支持向量机、随机森林等) 能捕捉时态数据中的复杂非线性关系。这些模型通过特征工程和序列化处理, 可以提升其数据分析的精确性。然而, 该类方法虽能处理更多元的模式, 但由于其主要依靠特征工程, 且需要基于专家领域先验知识手动设计特征, 因此存在无法完全自动化特征提取和无法充分挖掘数据中复杂的潜在模式等问题。

随着计算硬件的算力提升和数据集规模的不断增大, 深度学习逐渐兴起, 尤其是循环神经网络、长短期记忆网络、图神经网络和 Transformer 等模型在时态分析中的广泛应用, 使得时态数据分析进入了新阶段。这类模型能自动从数据中学习时态特征, 并捕捉远距离的时间依赖性, 极大提高了建模能力^[9-10]。然而, 深度学习算法具有较强的数据依赖性, 需要大量标注数据和计算资源, 且只有事先对模型进行训练才可使用, 因此该类方法存在泛化能力较弱和缺乏少样本、零样本推理能力等缺点, 这引起了基于大模型方法的时态数据分析算法的兴起。

由于传统深度神经网络需要收集大量特定领域的时态数据用于训练, 且对数据的要求较高, 因此模型的训练时间较长、成本较高, 且训练好的模型只能应用于特定领域。受大模型具有较强的泛化能力启发, 为增强各种时态模型的泛化能力和零样本推理能力^[11], 许多研究人员也将大模型运用到时态数据分析中, 构建了各种基于大模型的时态分析模型。与传统深度神经网络模型相比, 基于大模型的时态分析模型具有更强的泛化性, 预训练的模型能在各种时态分析任务上取得较好性能。

基于大模型的时态数据分析算法已经被广泛应用到各种实际任务中, 如金融数据分析^[12-13]、医疗数据分析^[14-15]、交通数据分析^[16-17]和天气气候分析^[18-19]等。此外, 已经有许多通用的分析算法能应用于各种特定领域的时态数据分析任务

中^[20]。虽然基于大模型的时态数据分析算法已经得到广泛应用, 且取得不错的效果, 但具备时态数据分析能力的大模型的发展仍处于初级阶段^[21], 也有一些工作表明, 将大模型应用于时态数据分析并非是一个很好的选择^[22-23], 各类方法在实际应用中也还存在着一些问题。

(1) 预训练的 LLM 并不天然具备理解时态数据的能力, 不能充分发挥其强大的泛化性能。在输入阶段, LLM 处理的是离散词素 (token), 而时间序列数据本质上是连续的, 这导致时态数据无法直接编辑或在自然语言中无损描述, 进而导致 LLM 不能很好地理解输入的时态数据, 推理效果不佳。研究表明, LLM 虽然有潜力完成时态数据分析任务, 但其只能完成简单的推理任务, 复杂推理能力差^[24]。

(2) 时态数据结构复杂, 难以与 LLM 的输入对齐。时态数据 (特别是时空数据) 的数据结构并不固定。这导致难以找到一种最优的嵌入方法使得 LLM 能输入各种时态数据。此外, 不同的输入数据可能具有不同的最优嵌入方法。嵌入方法直接关系到 LLM 的理解与推理能力。

(3) 重新训练模型难度高, 计算成本大。为解决上述问题, 许多工作开始研究基座模型, 将专为时态数据设计的模型替换掉 LLM, 但由于基座模型需要在时态数据上对大模型进行重新训练, 而大模型参数量大, 训练困难, 且难以获取庞大的数据集以训练大模型, 因此构建基座模型难以推广。

(4) 难以从自然语言输出中提取分析结果, 提取过程会损失分析性能。由于 LLM 是为自然语言设计的, 且为使输出更符合人类的意图, 对 LLM 进行了对齐技术, 导致其在时态数据分析上的输出结果表现不佳。相关工作表明, 基于人类反馈的强化学习会影响 LLM 在时序数据分析上的性能^[25]。

针对上述问题, 尽管现在的方法在大模型推

理流程的各阶段都有一定的改进，但这些问题仍没有被很好地解决。

为了让读者更全面地了解该领域的最新学术动态与进展，本文系统地分类整理了相关工作，并阐述了该领域仍面临的问题与未来发展方向。

2 相关概念

2.1 大模型

大模型指在大量数据上进行预训练，并能在多种下游任务中表现良好的深度学习模型，通常具有非常大的参数规模，能处理复杂的任务，并生成高质量的输出。大模型的核心思想如下：通过在大规模数据集上进行预训练，捕捉广泛的语言模式或其他领域的知识，从而在面对特定任务时仅需少量调整或微调。近年来，随着大数据集与算力的快速发展，由于大模型具有强大的泛化能力与高可扩展性，因此越来越多的大模型被应用到各领域。例如，在计算机视觉领域，大视觉模型被证明是可扩展的，可运用于各种下游任务^[26]；在NLP领域，GPT-4^[27]和Llama 2^[28]等LLM凭借着极强的文本理解与生成能力，已被应用于生活中的各方面。

2.2 时序数据分析

时间序列数据是按时间顺序排列的一组数据点，每个数据点通常与一个特定的时间戳相关联，通常表示为 $\mathbf{X}=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\} \in \mathbb{R}^{L \times D}$ ，其中， $\mathbf{x}_t \in \mathbb{R}^D$ 为在时间 t 的数据点， D 为序列维度， L 为序列长度。

时间序列分析可分为以下几类任务。

(1) 时序预测。时序预测指通过历史时间序列数据，预测未来某一时刻或一段时间内的数值或趋势。给定输入序列 $\mathbf{X}=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\}$ ，时序预测的任务是预测出序列 H 步的未来值 $\mathbf{X}_H=\{\mathbf{x}_{L+1}, \mathbf{x}_{L+2}, \dots, \mathbf{x}_{L+H}\}$ 。

(2) 时序插补。时序插补指在时间序列数据

中填补缺失值。给定输入序列 $\mathbf{X}=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\}$ ，缺失值标记为 $\mathbf{M}=\{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_L\}$ ，当 $\mathbf{x}_{t,d}$ 缺失时 $\mathbf{m}_{t,d}$ 为0，否则为1，时序插补的任务是给出一个填补序列 $\hat{\mathbf{X}}$ ，使得最终完整序列为 $\mathbf{X}_{\text{imputed}}=\mathbf{X} \odot \mathbf{M} + \hat{\mathbf{X}} \odot (\mathbf{1} - \mathbf{M})$ 。

(3) 时序分类。时序分类指基于时间序列数据进行分类。时序分类的任务是给出一个分类器 $f(\cdot)$ ，将输入的序列 \mathbf{X} 映射到 K 个类别中。

(4) 异常检测。时序异常检测旨在识别给定时间序列数据中的异常或偏离正常模式的点。给定输入序列 $\mathbf{X}=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\}$ ，异常检测的任务是给出每个时间点的异常值得分 $\mathbf{s}=\{s_1, s_2, \dots, s_L\}$ ，以判定某个时间点或输入序列的子序列是否为异常状况。

2.3 时空数据分析

时空数据指同时包含时间和空间信息的数据，通常用于描述事物或现象在不同时间和空间维度上的变化。时空数据的形式为 $\mathbf{Z}=\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_L\}$ ，其中，第 i 个数据点为 $\mathbf{z}_i=(t_i, \mathbf{s}_i)$ ， t_i 为该点的时间数据， \mathbf{s}_i 为该点的空间数据。不同空间数据的表示方式存在显著差异，因此无固定表示形式，可以是坐标数据、图数据等。时空数据分析的下游任务通常是特定场景下的应用，主要包括时空数据预测、时空异常检测和时空数据生成等。

2.4 基于大模型的时态分析模型

基于大模型的时态分析模型是应用了大规模预训练模型的时态数据分析模型。以核心模型在设计时是否专门考虑时态数据作为分类标准，基于大模型的时态分析模型可以分为基于LLM的时态分析模型和时态基座模型，本文将分别阐述这两类模型的相关工作。

本文设计了一个统一的分类框架，对现有研究依据其针对的优化方向进行分类。具体而言，相关工作的创新点往往集中体现在对其中某一特定模块的改进或优化上。大模型的推理流程通常

可划分为 4 个核心步骤: 输入、嵌入、模型 (架构)、输出。基于这一流程框架, 根据研究重点聚焦的模块差异, 可将现有研究成果系统性地划分为 4 个对应类别: 输入优化、嵌入优化、架构优化和输出利用, 整体的分类框架如图 1 所示。本文按照此分类方法对相关工作进行了总结, 并

对每类方法进行优缺点评估与适用场景分析, 阐述不同方法之间的异同与发展趋势, 为在不同任务条件下选择将大模型嵌入时态数据分析模型的方式提供参考。对于研究人员与用户来说, 各类时态数据分析算法具有不同的应用价值和优势。

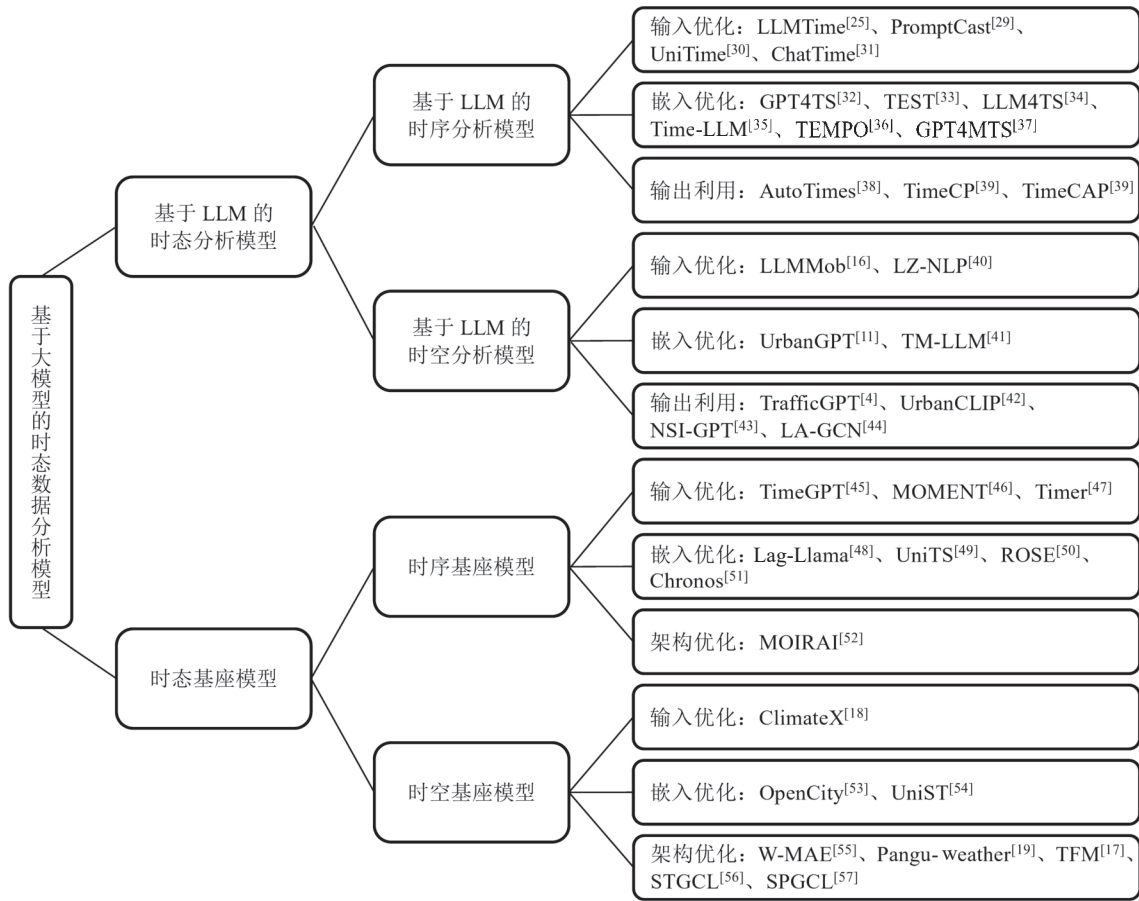


图 1 模型分类

Fig. 1 Models classification

(1) 输入优化。这类基于大模型的时态数据分析算法往往以提示 (prompt) 为基础或大量数据集为基础, 允许用户无须编写代码即可使用模型完成时态数据分析任务, 同时允许用户调用大模型的应用程序接口, 在无须本地部署模型的情况下实现分析功能, 有效降低了计算资源成本。

(2) 嵌入优化。这类基于大模型的时态数据

分析算法利用高效的方法将时态数据编码成嵌入输入, 泛化性能较强, 在多种下游任务应用场景中可以实现零样本、少样本推理, 允许用户不进行微调的情况下也能取得较好的分析性能。

(3) 架构优化。这类基于大模型的时态数据分析算法一般是针对特定应用领域构建的模型, 当用户根据任务所在领域选择这类模型时, 其性

能往往优于通用模型。

(4) 输出利用。这类基于大模型的时态数据分析算法将大模型作为工具，辅助用户在其他应用中取得更好的性能。此类方法同样无须用户在本地部署大模型，计算成本较低。

3 基于大语言模型的时态分析模型

基于 LLM 的时态分析大模型借助现有的 LLM，如 GPT-4^[27]和 Llama 2^[28]等，将 LLM 应用到时态分析任务中，以发挥 LLM 的泛化能力。时态数据分为时序数据和时空数据，时序数据有固定格式，数据处理简单，而时空数据无固定格式，不同数据之间的格式不同。两种数据与 LLM 结合的方法差异大，因此根据模型处理数据的形式，可分为时序分析模型和时空分析模型。基于 LLM 的时态数据分析模型的核心模型、适用领域及任务和主要特点如表 1 所示。

大模型的工作流程分为 4 个步骤：输入、嵌入、模型(推理)、输出。基于 LLM 的时序分析模型除了模型固定为 LLM 外，相关工作在其他 3 个步骤均有针对性的优化或利用。各类基于 LLM 的时态分析模型机制如图 2 所示。输入优化类模型主要将时态数据直接与 LLM 的文本输入对齐，利用 LLM 的能力实现时态数据分析；嵌入优化类模型主要为时态数据专门设置一个分词器，将时态数据直接与 LLM 的嵌入空间对齐；输出利用则将 LLM 作为一个辅助工具，将 LLM 与其他算法结合进行数据分析。

3.1 基于大语言模型的时序分析模型

LLM 的泛化能力很强，在各领域的任务中都表现出不错的性能。时序数据是一种结构化的数据，与自然语言具有相似的序列特性，因此将时序分析任务与 LLM 结合较简单。相关工作主要集中在通过提示工程、重新构建 LLM 的某些层(如嵌入层)和使用 LLM 代理等方法实现

LLM 与时态数据的结合。本节将相关工作分为模型输入优化、模型嵌入优化和模型输出利用。

3.1.1 模型输入优化

由于 LLM 在大量语言文本数据上训练，因此，LLM 并不具备理解时序数据中时序模式的能力。为使 LLM 能从文本中捕获这种时序模式，并应用到时序分析的下游任务中，则需要将时序数据与 LLM 的输入对齐。模型输入优化的方法主要分为两类：一类方法直接将时间序列转化为文本表示，并输入 LLM，以完成下游任务；另一类方法则为时间序列专门设计一个分词器，将时间序列转换为 token 再输入 LLM 中。

Gruver 等^[25]提出的 LLMTime 模型和 Xue 等^[29]提出的 PromptCast 模型是直接与时序数据以文本表示的典型方法。LLMTime 模型直接将时间序列数据编码为数字字符串，并将其转化为文本补全任务。该方法去除了时序数据的小数点外，还增加了空格，使 LLM 的分词器为每个数字分配一个 token，允许 LLM 进行零样本时间序列预测。而 PromptCast 模型则先将时序数据嵌入文本模板，再输入 LLM，预测范式不同于训练神经网络，即将时序信息的输入转换为 prompt，通过与 LLM 问答的方式完成时序预测任务。该模型设计了一系列 prompt 模板，允许将时序数据嵌入模板中，并输入 LLM，以获取分析结果。PromptCast 模型提供了一个“少代码”的视角，使得非计算机领域工作者也能使用 LLM 进行时序预测。这类直接将时序数据转化为文本输入 LLM 的方法无须对大模型进行训练，无须在下游任务上进行微调即可完成时序分析等任务。然而，由于时序数据无法在自然语言中进行无损描述，因此这类方法不能将时序信息完整地输入模型中，存在分析性能损耗的问题。

为时序数据设计分词器能有效减少时序数据转换为文本表示时的信息损失。Liu 等^[30]提出了 UniTime 模型，该模型有一个时序分词器，可将

表 1 基于 LLM 的时态数据分析模型

Table 1 Temporal data analysis models based on LLM

类型	优化方向	模型	年份	核心模型	领域	适用任务	主要特点
基于 LLM 的时序分析模型	输入优化	LLMTime ^[25]	2023	GPT3/Llama2-70B	通用	预测	直接将时间序列数据编码为数字字符串
		PromptCast ^[29]	2023	通用 (10 个 LLM)	通用	预测	将时序数据嵌入文本模板再输入 LLM
		UniTime ^[30]	2024	GPT2	通用	预测	设计了一个时序分词器, 并采用 Transformer 模块整合文本和时序 token
		ChatTime ^[31]	2025	Llama-2-7B	通用	预测	将时间序列视作一门外语生成嵌入, 并使用现有的预训练 LLM 进行微调
	嵌入优化	GPT4TS ^[32]	2023	GPT2	通用	预测/分类/补全/异常检测	重新设置了嵌入层, 将时间序列数据映射为 token, 并输入到 NLP 模型中
		TEST ^[33]	2023	通用 (4 个 LLM)	通用	预测/分类	通过时序实例、特性和文本原型对齐, 构建了一种时序 token 范式
		LLM4TS ^[34]	2023 (2025 出版)	GPT2	通用	预测	将时间序列归一化并切成补丁输入 LLM 中
		Time-LLM ^[35]	2023	Llama-7B	通用	预测	在时间序列切成补丁后引入了补丁重编码模块
		TEMPO ^[36]	2023	GPT2	通用	预测	提出了一种基于特征提取的软提示构建方法
		GPT4MTS ^[37]	2024	GPT2	通用	预测	针对时序数据和文本数据分别设计嵌入层
输出利用	AutoTimes ^[38]	2024	GPT-2/OPT/Llama	通用	预测	将时间戳输入 LLM, 以形成位置编码	
	TimeCP ^[39]	2025	通用 (默认 GPT-4+BERT)	通用	预测	引入了两个独立的 LLM 代理	
	TimeCAP ^[39]	2025	通用 (默认 GPT-4+BERT)	通用	预测	在 TimeCP 模型的基础上, 增加了一个多模态编码器	
	LLMMob ^[16]	2023	GPT3.5	交通	预测	将用户移动数据结合预设的 prompt 输入语言模型	
基于 LLM 的时空分析模型	输入优化	LZ-NLP ^[40]	2025	通用 (15 个 LLM)	交通	预测	将轨迹历史数据与上下文插入比 LLMMob 更简单的 prompt 模板
		UrbanGPT ^[11]	2024	Vicuna-7B	交通	预测	集成了一个多层次时间卷积网络的时空编码器
嵌入优化	TM-LLM ^[41]	2024	GPT2	交通	预测	为船只轨迹数据重新设计嵌入对齐	
	TrafficGPT ^[4]	2024	通用 (4 个 LLM)	交通	通用	使用 LLM 构造交通领域辅助分析工具	
	UrbanCLIP ^[42]	2024	Llama-Adapter V2	城市计算	预测	使用 LLM 构造图文对比, 基于图文对比对训练 CLIP 模型	
	NSI-GPT ^[43]	2023	ChatGPT	金融	预测	使用 LLM 的图推理能力增强神经网络	
	LA-GCN ^[44]	2025	BERT	动作识别	特征提取	将 LLM 用于时空特征提取	

时序数据转换成 token, 采用独特的语言-时间序列的 Transformer 模块整合文本和时序 token, 以解决时间序列域混淆的问题。该模型以 GPT2 模型^[58]为主干重新设计了语言模型, 可以显式地接受传入的时间序列域信息, 将来自不同输入空间的时间序列对齐到语言模型的共同隐空间, 实现

跨域泛化。为进一步兼容先进 LLM 的训练范式, ChatTime 模型^[31]将时间序列视作一门外语生成嵌入, 并使用现有的预训练 LLM, 对 LLM 进行继续预训练和指令微调, 以适配时间序列。该模型具有语言和时序的双模态能力, 除了能完成传统的时序分析任务外, 还具有时序问答等功

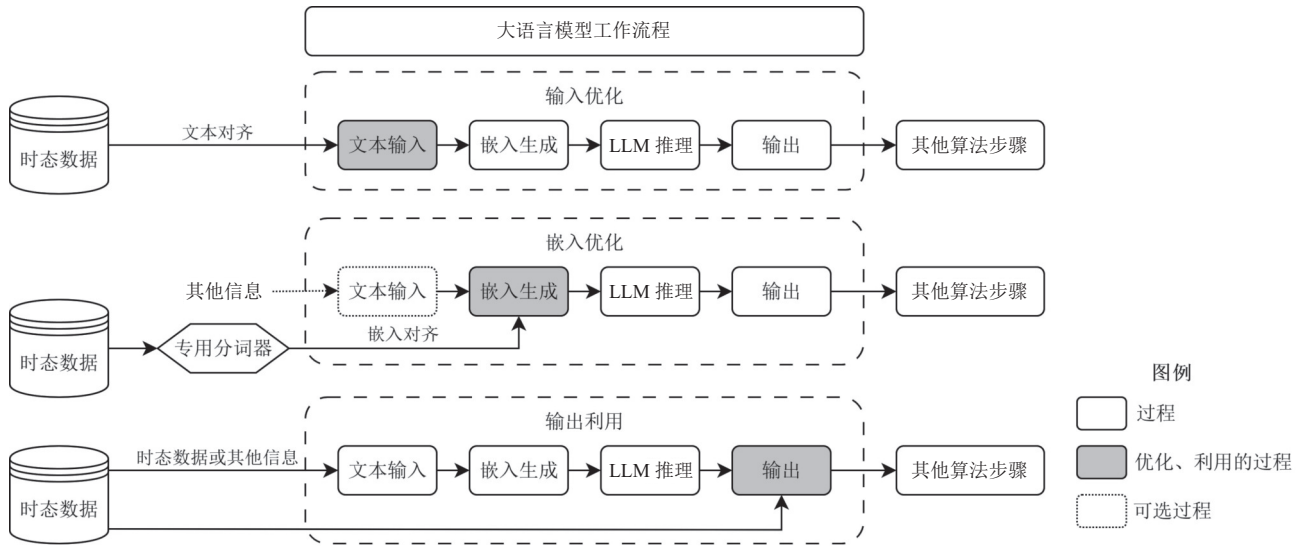


图 2 基于 LLM 的时态分析模型机制示意图

Fig. 2 Schematic diagram of mechanism of temporal analysis models based on LLM

能。与直接将时序数据输入 LLM 的方法相比，该类方法能在一定程度上减少输入过程的信息损耗；然而，由于 LLM 并不具备理解时序 token 的能力，因此该类方法需要对 LLM 进行进一步训练，计算成本较高。

3.1.2 模型嵌入优化

由于自然语言不具备无损表示时间序列数据的能力，因此将时序数据直接转换成嵌入并输入 LLM 是一种有效的对齐方法。Zhou 等^[32]提出的 GPT4TS 模型是首个将时序数据直接转换为嵌入输入 LLM 的工作，是后续许多工作的基线模型。为解决缺乏大量训练数据的困难，该模型使用了预训练的 NLP 模型，并重新设置了嵌入层，将时间序列数据映射为 token，并输入到 NLP 模型中。该模型提出了一种冻结的预训练 Transformer，其中 Transformer 块中的自注意力层和前馈层被冻结，只有嵌入层、归一化层和输出层需要训练。该工作证明了 Transformer 模块中的自注意力层的机制类似于主成分分析，从原理上解释了将 LLM 应用于时序分析的迁移学习是有效的，为后续工作提供了重要的理论支撑。在 GPT4TS 模型提出后，后续的相关工作主要分

为两类。

一类工作进一步优化时序嵌入的性能，在 GPT4TS 模型证明了将 LLM 迁移至时间序列的有效性的基础上，Sun 等^[33]证明了软提示和监督微调之间的等价性，并提出了 TEST 模型。将多变量时间序列转换为文本输入 LLM 会变成单变量序列，丢失大量时序特征，因此 TEST 模型应用了一种全新的嵌入方法。TEST 模型未直接将时序数据作为文本输入 LLM 中，而是通过将时序实例、属性与文本原型对齐，构建了一种时序 token 范式，跳过了原 LLM 的嵌入层，直接编码嵌入到 LLM 中，并训练生成一种软提示，让 LLM 接受时序输入。该方法的性能与 GPT4TS 模型相似，在不同数据集上各有优劣。然而，这些方法都没有很好地解决多尺度时序分析的问题，因此，Chang 等^[34]提出了 LLM4TS 模型，旨在解决 LLM 与时序数据难以对齐和 LLM 难以理解时序数据的多尺度信息这两个问题。针对对齐任务，该模型提出了两步解决方法，通过与 GPT4TS^[32]类似的自回归目标法，将时间序列归一化，并切成补丁输入 LLM 中，使 LLM 与时序数据特征对齐，再按照相应的时序分析任务进

行训练, 以获得良好的泛化能力。针对激活 LLM 理解多尺度信息的任务, 该模型提出一种两级融合策略, 将时序数据的多尺度信息嵌入补丁中, 帮助激活 LLM 进一步理解时序数据的能力。由于 LLM4TS 模型具有额外的时间序列对齐和多尺度时间信息集成机制, 因此该模型的预测性能全面超过了 GPT4TS 模型。此类优化时序嵌入性能的方法直接将时序嵌入对齐 LLM 输入, 无须考虑文本输入的对齐, 但导致在文本上预训练的 LLM 损失了一定推理能力。

另一类工作则将时序数据与文本输入对齐, 选择将 prompt 的嵌入拼接至时序嵌入之前, 使模型具有双模态能力。在 GPT4TS^[32]模型对齐方法的基础上, Jin 等^[35]提出了 Time-LLM 模型, 在时间序列切分成补丁后引入了补丁重编码模块, 允许模型自适应地选择相关的输入信息, 而不离开语言模型预先训练的空间。这种方法可使预训练模型无须更新参数即可产生所需的输出。此外, 该方法还引入了“提示前缀”, 在输入处理好的时序数据前, 采用一个设定好的提示模板, 用额外的上下文信息丰富输入时间序列, 并以自然语言的形式为 LLM 提供任务指令。同时期提出的还有 TEMPO 模型^[36], 该模型采用了一种基于特征提取的软提示构建方法, 通过局部加权回归的方法将时间序列输入分解为趋势、季节性和残差 3 个分量, 并将各分量映射到隐空间中。时间序列通过经典的“归一化-切分补丁”的方法形成嵌入, 与隐空间中的特征分量一同输入到 LLM 中, 以实现时序预测。然而, 这些方法所输入的 prompt 是固定的模板, 跨模态推理时的灵活性较差。因此, Jia 等^[37]提出了 GPT4MTS 模型, 该模型采用基于 prompt 的 LLM 框架, 可以同时利用数值数据和文本信息进行多模态时间序列预测。该模型支持自定义输入 prompt 以辅助时序分析, 且针对时序数值数据和文本信息分别设计嵌入层, 并使用预训练好的 GPT-2 模型作

为 LLM 框架。此类将文本嵌入与时序嵌入对齐的方法具有跨模态推理的能力, 零样本、少样本推理的性能更好。

针对嵌入对齐的模型需要微调训练 LLM 的一些神经网络层, 对计算资源有一定需求。由于时序数据以大语言模型能理解的格式进入大模型, 因此预测效果较好。相关研究表明, 通过将时间序列转换为补丁微调预训练的 LLM, 并进行时间序列预测, 是一种前景较好的策略^[59]。

3.1.3 模型输出利用

由于 LLM 的上下文理解能力很强, 因此将 LLM 视作工具, 并将 LLM 输出与其他传统时序分析模型结合, 能以较低的计算成本将 LLM 应用到时序分析中。Liu 等^[38]提出了 AutoTimes 模型, 利用 LLM 的自回归特性实现时序预测。该模型将时间戳输入 LLM, 以形成位置编码, 利用该编码中的时间信息对齐同一时刻不同变量的时序片段。该位置编码与时间序列分段输入 LLM, 根据 LLM 输出的 token, 用自回归的方法还原并输出时间序列段, 以进行时序预测。随后, Lee 等^[39]提出的 TimeCP 模型创造性地引入了两个独立的 LLM 代理, 即上下文 LLM 代理和预测 LLM 代理, 利用 LLM 的上下文理解能力进行时间序列事件预测。该模型首先使用上下文 LLM 代理, 输入相应的 prompt 将时间序列转化为一段摘要, 摘要包含了原始时间序列数据之外的相关上下文理解。随后, 预测 LLM 代理根据摘要对时间序列做出预测。实验表明, 与单 LLM 代理的方法 (如 PromptCast 模型^[29]) 相比, 该模型的性能更好。TimeCAP 模型^[39]则在 TimeCP 模型^[39]的基础上, 增加了一个多模态编码器, 该编码器引入了输入增强和 prompt 增强, 同时输出预测结果和增强的 prompt, 与 LLM 代理相互补充。TimeCAP 模型将采样相关的文本摘要作为提示增强, 帮助预测 LLM 代理做出预测。最终的预测结果由多模态编码器的预测结果和预测

LLM 代理的预测结果融合而成。

此类模型具有无须训练大模型的特点，且允许调用大模型的 API，在本地算力较低的情况下也能使用大数量的 LLM 取得更好的分析性能。然而由于此类模型完全冻结 LLM，因此可能存在灵活性差、无法针对特定任务调整 LLM 等缺点。

3.2 基于大语言模型的时空分析模型

与时序分析不同，时空数据无统一的数据格式，因此难以与 LLM 对齐。即使能在某些数据上与 LLM 对齐，相关的模型也不能广泛地适用于时空数据。因此，基于 LLM 的时空分析大模型的相关研究较少，主要以利用 LLM 为主。与时序分析模型类似，基于 LLM 的时空分析模型的相关工作也分为模型输入优化、模型嵌入优化和模型输出利用。

3.2.1 模型输入优化

由于时空数据结构复杂，难以在自然语言中表述。因此，为将时空数据插入文本输入中，Wang 等^[16]提出了 LLMMob 模型，该工作探讨了 LLM 在人类移动预测任务中的潜力。该工作的任务为预测移动轨迹，其方法为将用户移动数据拆分为历史地点、当前停留地点、期望停留时间等信息，结合预设的 prompt 生成文本输入，并输入 LLM，LLM 输出预测的未来位置和相关原因。该模型成功将 LLM 迁移至时空预测任务。为构建更简单、高效的轨迹预测模型，Beneduce 等^[40]提出了比 LLMMob 模型^[16]更简化的 LZ-NLP (language-based zero-shot next location predictors) 模型，将轨迹历史数据和上下文插入更简单的 prompt 模板，由 LLM 进行轨迹预测的零样本推理。

由于时空数据比时序数据更复杂，因此完整地将时空数据以自然语言表述更困难，同时，LLM 并不天然地具有时空推理能力。由于这些问题未被很好地解决，因此输入优化的基于 LLM 的时空分析模型性能较低，相关工作较少。

3.2.2 模型嵌入优化

LLM 难以从自然语言中理解时空数据，因此，在时空数据输入 LLM 前，设计时空特征提取器，并将时空数据转化为嵌入再输入 LLM 成为一种主流方法。在城市数据分析领域中，为解决 LLM 难以理解时空数据中固有时间演化模式的问题，Li 等^[11]提出了 UrbanGPT 模型。该模型具有一个由多层次时间卷积网络构成的时空编码器，以不同分辨率捕获复杂的时间依赖关系。编码后的时空嵌入通过一个由线性层构成的轻量化对齐模块与 LLM 的输入对齐，使得该模型在不同的时空场景中具有较强的泛化能力。在航行轨迹预测领域中，Nguyen 等^[41]提出了 TM-LLM 模型，该模型直接丢弃了原始的 LLM 文本嵌入层，为船只轨迹数据重新设计嵌入对齐。该方法利用一维卷积神经网络提取局部时空特征，利用双向的长短期记忆网络提取长期依赖特征。提取的两种特征经过融合后与 LLM 的嵌入对齐，并进行微调。LLM 的输出经过回归卷积层得到对船只的轨迹预测。通过特征提取方法将时空数据转换为嵌入能激活 LLM 对时空数据的理解能力，比输入优化的模型的分析性能更好。

3.2.3 模型输出利用

由于时空数据难以与 LLM 对齐，因此模型输出利用这一类方法的许多初期研究是将 LLM 作为其他模型的辅助工具，利用 LLM 的输出协助其他模型完成分析任务。输出利用的工作一部分是将 AI 代理应用于时空分析任务中，而另一部分则是将 LLM 作为特征提取器，帮助下游模型在时空数据集上对各种模态的数据进行特征提取。

使用 AI 代理的典型模型是 TrafficGPT 模型^[4]，其使用 LLM 构造交通领域辅助分析工具，其中 LLM 作为任务调度和评估的组件，理解用户意图，调用专用小模型。这使得 LLM 能接受广泛格式的输入，同时也能提供专用基础模型精确的分析结果。该模型赋予 LLM 查看、分析和

处理交通数据的能力, 为城市交通系统的管理提供有效的决策支持。

由于 LLM 本身具备强大的多模态分析能力, 因此将 LLM 作为特征提取器可以直接从多模态数据中进行特征提取。在图片模态, UrbanCLIP 模型^[42]使用 LLM 为航拍图片生成高质量文本描述, LLM 作为图生文的模型构造图文对比对, 基于图文对比对训练 CLIP 模型, 从而生成文本描述。该文本描述将与轨迹、坐标和路网等信息一同进行对比预训练, 以增强模型在城市视觉分析下游任务中(如人口预测、污染预测等)的表现。在文本模态, Chen 等^[43]提出了 NSI-GPT (network structure inference via ChatGPT) 模型, 该模型利用 LLM 的图推理能力增强图神经网络。该方法将金融领域的新闻、公司信息和 prompt 输入 LLM, 并生成图信息, 该信息将进入后续的图神经网络, 以构建“受影响的公司”图关系, 并用于后续股票价格的分析。在图模态, Xu 等^[44]提出的 LA-GCN 模型将 LLM 用于时空特征提取, 该模型针对的是人类骨架动作序列输入, 并利用骨架运动模式识别行为。该模型将骨架动作序列输入 LLM, 并利用其知识辅助所使用的图卷积网络进行动作识别。这 3 项处理不同模态数据的工作均表明可以使用 LLM 作为特征提取器, 结合其先验知识, 帮助专用模型提高性能。

LLM 的输出利用探索主要从 AI 代理和特征提取两方面展开。这些探索既规避了直接对齐时空数据与 LLM 的困难, 又有效发挥了 LLM 的语义解析、知识推理和跨模态对齐优势, 为时空数据分析系统提供了高泛化性的解决方案。

4 时态基座模型

时态基座模型也称为时态基础模型, 是通过大规模数据预训练得到的专门处理时态数据的预

训练模型, 具备强大的特征提取和泛化能力, 通常基于海量数据训练, 可通过微调适配多种下游任务。相关工作在构建基座模型时, 时序数据与时空数据的数据处理方法差异极大, 因此本节将相关工作分为时序基座模型和时空基座模型两类。

由于基座模型是为时态数据专门设计的, 其输出本身即服务于时态分析任务, 因此按照大模型推理的 4 个流程: 输入、嵌入、模型(架构)、输出, 现有研究主要集中在除输出流程外的其他 3 个维度的优化, 各类模型的机制如图 3 所示。输入优化模型主要采用大规模时态数据集对 LLM 进行训练, 以得到时态数据专用的分析模型; 嵌入优化模型则主要设计专用的特征提取器, 再将生成的嵌入输入核心模型(一般为 Transformer)进行推理; 架构优化模型主要是针对时态数据的特点设计的一种专门的推理架构。

4.1 时序基座模型

时序基座模型采用大模型的基本思路, 面向时序数据构建专用的基座模型, 拥有专用模型结构与训练方法。该模型面向标准时间序列数据, 即已在时间步上对齐的单元或多元时间序列数据。受 LLM 在跨领域任务上取得显著突破的影响, 时序基座模型的相关工作主要通过大规模时序数据训练构建基座模型。时序基座模型可以分为模型输入优化、模型嵌入优化和模型架构优化 3 个类别, 各模型的参数量、数据量、适用领域及任务和主要特点如表 2 所示。

4.1.1 模型输入优化

一些工作通过在大模型的输入阶段赋予大模型理解时间序列分析的能力, 以构建基座模型, 使模型在各种时序分析下游任务中能进行零样本、少样本推理。时序基座模型面临的一大困难是难以获取庞大的数据集来训练大模型, 而 Garza 等^[45]提出的 TimeGPT 模型则完全参考

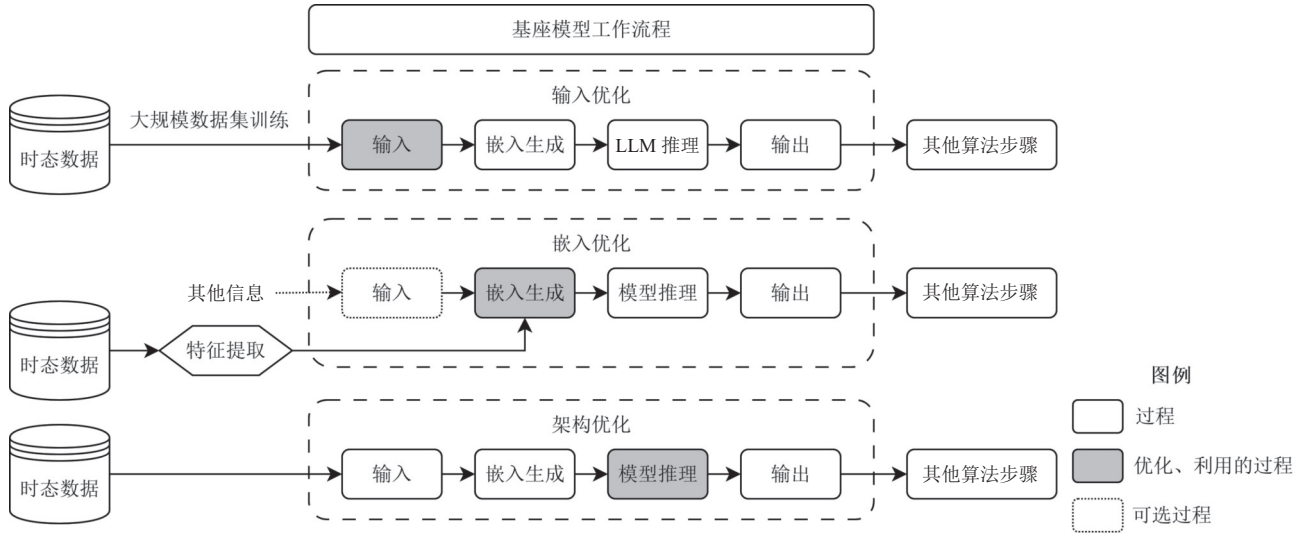


图3 时态基座模型机制示意图

Fig. 3 Schematic diagram of mechanism of temporal foundation models

表2 时序基座模型

Table 2 Time series foundation models

优化方向	模型	年份	参数量	数据量	领域	适用任务	特点
输入优化	TimeGPT ^[45]	2023	未知	100 B	通用	通用 (预测/分类/补全/异常检测)	在不同领域的数据集上进行了广泛的训练
	MOMENT ^[46]	2024	40 M, 125 M, 385 M	1.13 B	通用	通用 (预测/分类/补全/异常检测)	编制了一个庞大且多样化的时间序列数据集, 实现时序基座模型的预训练
	Timer ^[47]	2024	29 M, 50 M, 67 M	1 B	通用	通用 (预测/补全/异常检测)	将异构时间序列统一为单序列格式
嵌入优化	Lag-Llama ^[48]	2023	200 M	352 M	通用	预测	使用“滞后特征”捕捉时间序列数据的历史依赖性
	UniTS ^[49]	2024	未知	35 M	通用	通用 (预测/补全/异常检测)	使用 prompt 框架将各种任务转换为统一的 token 表示
	ROSE ^[50]	2024	7.4 M	0.89 B	通用	预测	使用补丁 token 和寄存器 token, 增强泛化能力
	Chronos ^[51]	2024	8 M, 46 M, 201 M, 710 M	84 B	通用	预测	借用现有的 LLM, 并重新设计嵌入, 再对 LLM 进行微调
架构优化	MOIRAI ^[52]	2024	14 M, 91 M, 311 M	27 B	通用	预测	提出基于编码器的掩码通用时序预测 Transformer 架构

GPT 模型的架构, 在不同领域的数据集上进行广泛的训练, TimeGPT 的训练样本超过了 1 000 亿个, 均来自开源数据集, 其领域包括金融、经济、人口统计学、医疗、天气、传感器、能源、网络流量、销售、运输等。该方法通过大规模训练提升了模型的推理能力, 然而, 该方法为闭源模型, 大型时序数据集仍然缺乏。因此, Goswami

等^[46]针对大型公共时间序列数据集缺少和时序数据集多样性缺乏等问题, 编制了一个庞大且多样化的时间序列数据集, 称为时间序列堆, 以实现大规模的多数据集预训练。其使用掩码 Transformer 模型在该数据集上构建了 MOMENT 时序基座模型, 该模型可以在各种时序分析下游任务上微调并应用。同一时期, Liu 等^[47]则将异构时间序列

统一为单序列格式, 采用时序 Transformer 架构在 10 亿个时间点上进行训练, 形成 Timer 模型。这两个方法解决了该领域缺乏大规模数据集及其训练的开源模型的问题。相关工作表明, Timer 模型在某些特定的数据集上是最优模型^[60]。这些模型取得了不错的零样本推理能力, 然而, 其缺陷是数据集过于庞大, 计算成本极高, 其应用推广可能会受到限制。

4.1.2 模型嵌入优化

通过设计高效提取时序特征的专用架构形成时序嵌入是降低重新训练一个基座模型的计算成本的可行方法。将时序数据信息与额外信息一同生成嵌入是构建嵌入优化基座模型的主要方法。Rasul 等^[48]提出了 Lag-Llama 模型, 该模型使用“滞后特征”捕捉时间序列数据的历史依赖性。滞后特征指过去时间点的数据能帮助模型理解和预测时间序列中的特定模式, 如趋势或季节性。通过观察历史值, Lag-Llama 模型可以学习到特定的模式, 并以概率形式预测未来的值。该方法虽然能有效捕获时序信息和进行预测, 但其嵌入方法相对固定, 且仅适用于时序预测任务。为统一各种时序分析的下游任务形式, Gao 等^[49]提出了一种通用的时序分析模型 UniTS, 该模型使用 prompt 框架将各种任务转换为统一的 token 表示, 从而创建适用于所有任务的通用规范, 该 token 表示法具有较强的灵活性, 支持嵌入不同的信息与设定分析任务, 且该模型能通过少量或无须额外训练即可适应新任务。然而, 这些方法仍然存在难以跨域泛化的问题。因此, Wang 等^[50]提出的 ROSE 模型则使用补丁 token 和寄存器 token, 增强 ROSE 模型从多训练域到不同目标域的自适应转移的泛化能力。补丁 token 将时间序列切分成小块, 而寄存器将来自多域数据集的特定领域的信息聚类到寄存器 token 中, 并在训练前将这些特定领域的信息存储在寄存器中, 与补丁 token 一同输入到 Transformer 模型中。

该模型是表现最佳的时序基座模型^[60]。

此外, 先借用现有的 LLM 重新设计嵌入, 再对 LLM 进行微调以构建时序基座模型, 能有效减少训练的计算成本。Ansari 等^[51]提出了 Chronos 模型, 该模型将时间序列按数值的 bin 离散化, 生成时序数据的“词表”, 随后通过掩码和交叉熵损失优化基座模型和词向量的嵌入。嵌入通过一个语言模型预测输出, 并映射回最终的数值。此类方法的优势是允许该模型成为 LLM 中一个“即插即用”的插件, 可以迅速套用到最新的先进 LLM 上, 并应用到时序分析任务中。

4.1.3 模型架构优化

针对时间序列专门设计的模型架构能高效完成时间序列推理, 减少计算成本。Woo 等^[52]提出了 MOIRAI 架构, 其全称为基于编码器的掩码通用时序预测 Transformer 架构。该架构类似于普通的 Transformer 架构, 是一个通用模块, 允许运用到各种时序分析任务中。该模型采用一种基于不重叠补丁的方法, 使用掩码编码器架构建模时间序列。该架构可以“扁平化”多变量时间序列, 将所有变量视为单个序列。随后, 通过多补丁大小的输入投影层, 将补丁投影到向量表示中, 并采用掩码表示需要预测的时间步。MOIRAI 架构提供了强大的零样本预测能力, 在跨领域和多样化频率的时间序列中均适用。

4.2 时空基座模型

时空基座模型是一种用于处理时空数据的基础模型, 能同时捕捉空间和时间的复杂依赖关系。由于时空数据没有固定的格式, 因此难以在时空数据分析领域构建一个通用模型。目前, 时空基座模型的相关工作较少^[61], 发展仍处于初级阶段。与时序基座模型类似, 时空基座模型也可以划分为模型输入优化、模型嵌入优化和模型架构优化。时空基座模型的适用领域及任务和主要特点如表 3 所示。

表 3 时空基座模型

Table 3 Spatio-temporal foundation models

优化方向	模型	年份	领域	适用任务	特点
输入优化	ClimateX ^[18]	2023	气象	通用	构建气象专用数据集，并用于训练 LLM
嵌入优化	OpenCity ^[53]	2024	交通	通用	构建了一种分布移位泛化的时空嵌入
	UniST ^[54]	2025	通用	预测	设计了一种通用的时空嵌入方法，适配各种城市时空预测的场景
架构优化	W-MAE ^[55]	2023	气象	通用	将视觉 Transformer 架构引入时空分析中
	Pangu-weather ^[19]	2022	气象	预测	设计了一种 Transformer 架构，并用于与地球的几何结构对齐
	TFM ^[17]	2023	交通	预测/模拟	利用基于生成图 Transformer 的图结构和动态图生成算法解决复杂的流量问题
	STGCL ^[56]	2022	通用	预测	将对比学习纳入时空图预测中
	SPGCL ^[57]	2022	通用	预测	通过最大化正负邻居之间的区分边际学习信息关系

4.2.1 模型输入优化

在不改动 LLM 架构的前提下，通过构建专用数据集训练 LLM 是构建基座模型的一种方法。ClimateX 数据集^[18]是一个全新的、精心设计的、具有专家标记的数据集。使用该数据集训练 LLM，可对人类专家对气候相关陈述的置信度进行分类。该类方法虽然具有一定的发展潜力，但难以构建大量相关的专用数据集和需要大量人力成本仍然是其面临的巨大挑战。

4.2.2 模型嵌入优化

根据时空数据设计嵌入方法可以让模型更好地提取数据中的时空特征。针对交通数据时空异质性的固有挑战，Li 等^[53]提出了 OpenCity 模型。该模型构建了一种分布移位泛化的时空嵌入，采用上下文归一化的方法实现零样本推理能力，使用补丁嵌入实现高效的长期预测。随后分别设计了时间上下文编码和空间上下文编码，以捕获不同交通场景的时空特征，最后采用基于 Transformer 架构的时空依赖建模实现时空预测。该模型在 21 个数据集上训练并评估，是时空基座模型中采用数据集领域较多的模型。而 Yuan 等^[54]提出的 UniST 模型则采用了一种通用的时空嵌入方法，允许模型适配各种城市时空预测的场景。该嵌入方法利用一维卷积网络和图卷积网络

处理时空数据，并采用随机、管道、块和时态 4 种掩码方法训练时空嵌入表示网络，并将生成的时空嵌入与知识导向 prompt 融合，使用 Transformer 架构实现通用的城市时空预测任务。该模型的嵌入方法更通用，与数据类型无关，场景适用性比 OpenCity 模型更好。

4.2.3 模型架构优化

根据相关的时空数据分析任务，优化相应的模型架构是提高分析性能的重要途径之一。相关工作主要分为两类，一类侧重于针对应用领域设计模型架构，另一类侧重于对模型设计高效的学习方法。

设计模型架构的工作大多基于 Transformer 框架进行修改。在天气预测领域，Man 等^[55]提出了 W-MAE 模型，该模型创新性地将视觉 Transformer 架构引入时空分析中。该方法使用掩码自动编码器和视觉 Transformer 构建了一个与任务无关的天气预测基座模型，可以被微调至各种特定任务。但该方法先将地球图展平成图片，再使用视觉 Transformer 进行分析，忽略了地球的三维空间信息。随后 Bi 等^[19]提出了 Pangu-weather 模型，该模型是一个基于深度学习的快速、准确的全球天气预报系统。根据该任务的时空图输入为时间和地球坐标，该工作提出了一个

3D 地球 Transformer, 专门设计了一种 Transformer 架构用于与地球的几何结构对齐, 该方法的分析维度更高, 性能更好。在交通领域, TFM^[17] 作为一个交通运输基础模型, 主要利用基于生成图 Transformer 的图结构和动态图生成算法, 克服传统交通模型中结构复杂性和模型准确性的挑战, 解决复杂的流量问题。针对交通数据, 该模型将交通参与者表示为节点, 交通关系表示为边, 以构成图数据。该工作表明 TFM 具备从真实数据中学习的能力, 在交通领域具有广泛应用。

设计学习方法的工作主要集中于对通用时空基座模型的理论基础研究。Liu 等^[56] 提出了一个通用的时空图预测基座模型 STGCL, 该工作首次将对比学习纳入时空图预测中。该模型主要针对数据量不足导致模型过拟合、缺乏泛化能力的问题, 构建出时空图数据的正负对比对, 采用对比学习的方法提取时空图的语义信息: 相关的时空图的语义表示向量相似度最大化; 不相关的时空图的语义表示向量相似度最小化。基于对比学习的自监督学习方法能有效解决模型泛化性能不足的问题。类似的工作还有 Li 等^[57] 提出的自步图对比学习方法 SPGCL。该方法通过最大化正邻居和负邻居之间的区分边际学习信息关系, 并生成一个具有自步策略的最优图。该模型是一种通用的时空学习模型, 着重强调监控位置之间的语义关系, 能在确定的邻接信息不可用的情况下, 灵活实现图结构的构建。这些工作对其他时空基座模型的设计具有启发性意义, 支持其他模型使用更高效的模型训练方法。

这些工作均表明, 针对专业领域中的时空图数据, 设计专用的模型架构, 并构建该领域的基座模型是一种有潜力的时空数据分析方法。此类方法仅针对特定领域的数据训练基座模型, 能通过减少数据类型来降低训练难度, 是具有发展前景的方法之一。

5 总结与展望

本文分类介绍了基于大模型的时态数据分析算法, 对各类方法下的模型进行了比较分析。研究表明, 与传统分析算法相比, 将 LLM 引入时态数据分析算法中能在一定程度上提升分析的准确率, 增强算法的泛化性和提高分析结果的可解释性^[62], 时态数据模态与 LLM 的语义对齐虽然仍面临重大挑战^[22], 但现有研究已充分验证了模态对齐的可行性与发展前景^[63-64]。然而, 与专用小模型相比, 大模型在零样本推理任务中展现出更强的泛化能力^[62], 但其分析精度可能不如在相关数据集上重新训练过的小型模型^[23]。

综合相关工作, 基于大模型的时态数据分析算法提供了一种全新的分析工具。该类算法允许用户不进行微调或在少量数据上微调即可将这些算法应用于相关任务上, 有效解决了数据缺乏、冷启动等问题。在各种基于大模型的时态数据分析算法相关工作中, Transformer 架构表现出强大的通用性, 能作为时态数据分析基座模型的核心架构, 针对特定任务设计的 Transformer 衍生模型则展现出更优的性能。但基于大模型的时态数据分析算法的发展仍呈现初级阶段特征, 主流方法仍聚焦于对现有 LLM 能力的直接调用, 针对时态数据特性进行深度模型微调的工作相对匮乏。

该领域未来的研究方向可能包括但不限于以下 3 个维度: 一是持续拓展 LLM 的时态数据理解能力边界, 可以通过更先进的训练方法提高 LLM 在文本数据上的时态数据理解能力, 也可以通过设计创新的模型架构提高 LLM 在理论层面上的时态数据理解能力, 后者对模型的性能提升较大; 二是研发适配时态数据特性的专用分词器, 将时态数据映射到 LLM 能理解的文本空间仍然是一个有价值的研究方向; 三是在尺度定律指导下通过扩展模型参数量释放大模型的潜在性能, 更大的参数量与训练数据量能提升大模型的

推理性能。这些研究方向将共同推动时态数据分析大模型从理论探索走向实际应用。

参 考 文 献

- [1] Yu XL, Chen Z, Ling Y, et al. Temporal data meets LLM—explainable financial time series forecasting [Z/OL]. arXiv Preprint, arXiv: 2306.11025, 2023.
- [2] Morid MA, Sheng ORL, Dunbar J. Time series prediction using deep learning methods in healthcare [J]. *ACM Transactions on Management Information Systems*, 2023, 14(1): 1-29.
- [3] Barrera-Animas AY, Oyedele LO, Bilal M, et al. Rainfall prediction: a comparative analysis of modern machine learning algorithms for time-series forecasting [J]. *Machine Learning with Applications*, 2022, 7(1): 100204.
- [4] Zhang SY, Fu DC, Liang WZ, et al. TrafficGPT: viewing, processing and interacting with traffic foundation models [J]. *Transport Policy*, 2024, 150(6): 95-105.
- [5] Lu RZ, Bai RC, Li RD, et al. A novel sequence-to-sequence-based deep learning model for multistep load forecasting [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 36(1): 638-652.
- [6] 陈佳雯, 陈金荣, 陈星, 等. 时空数据驱动的智能家居服务管控方法 [J]. *集成技术*, 2024, 13(4): 16-29.
Chen JW, Chen JR, Chen X, et al. A spatio-temporal data-driven control method for smart home service [J]. *Journal of Integration Technology*, 2024, 13(4): 16-29.
- [7] Jin M, Wen QS, Liang YX, et al. Large models for time series and spatio-temporal data: a survey and outlook [Z/OL]. arXiv Preprint, arXiv: 2310.10196, 2023.
- [8] Ahmed NK, Atiya AF, Gayar NE, et al. An empirical comparison of machine learning models for time series forecasting [J]. *Econometric Reviews*, 2010, 29(5-6): 594-621.
- [9] Wang SZ, Cao JN, Yu PS. Deep learning for spatio-temporal data mining: a survey [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2020, 34(8): 3681-3700.
- [10] Lim B, Zohren S. Time-series forecasting with deep learning: a survey [J]. *Philosophical Transactions of the Royal Society A*, 2021, 379(2194): 20200209.
- [11] Li ZH, Xia LH, Tang JB, et al. UrbanGPT: spatio-temporal large language models [C] // *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024: 5351-5362.
- [12] Zhang BY, Yang HY, Liu XY. Instruct-FinGPT: financial sentiment analysis by instruction tuning of general-purpose large language models [Z/OL]. arXiv Preprint, arXiv: 2306.12659, 2023.
- [13] Xie QQ, Han WG, Lai YZ, et al. The wall street neophyte: a zero-shot analysis of chatGPT over multimodal stock movement prediction challenges [Z/OL]. arXiv Preprint, arXiv: 2304.05351, 2023.
- [14] Li J, Liu C, Cheng SB, et al. Frozen language model helps ECG zero-shot learning [C] // *Proceedings of the Medical Imaging with Deep Learning*, 2024: 402-415.
- [15] Liu X, McDuff D, Kovacs G, et al. Large language models are few-shot health learners [Z/OL]. arXiv Preprint, arXiv: 2305.15525, 2023.
- [16] Wang XL, Fang M, Zeng ZC, et al. Where would i go next? Large language models as human mobility predictors [Z/OL]. arXiv Preprint, arXiv: 2308.15197, 2023.
- [17] Wang XH, Wang D, Chen L, et al. Building transportation foundation model via generative graph Transformer [C] // *Proceedings of the 2023 IEEE 26th International Conference on Intelligent Transportation Systems*, 2023: 6042-6047.
- [18] Lacombe R, Wu K, Dilworth E. ClimateX: do LLMs accurately assess human expert confidence in climate statements? [Z/OL]. arXiv Preprint, arXiv: 2311.17107, 2023.
- [19] Bi KF, Xie LX, Zhang HH, et al. Pangu-weather: a 3D high-resolution model for fast and accurate global weather forecast [Z/OL]. arXiv Preprint,

- arXiv: 2211.02556, 2022.
- [20] Jiang YS, Pan ZJ, Zhang XK, et al. Empowering time series analysis with large language models: a survey [Z/OL]. arXiv Preprint, arXiv: 2402.03182, 2024.
- [21] Jin M, Zhang YF, Chen W, et al. Position: what can large language models tell us about time series analysis [C] // Proceedings of the 41st International Conference on Machine Learning, 2024: 22260-22276.
- [22] Zheng LN, Dong C, Zhang WE, et al. Understanding why large language models can be ineffective in time series analysis: the impact of modality alignment [C] // Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V, 2025: 4026-4037.
- [23] Tan MT, Merrill MA, Gupta V, et al. Are language models actually useful for time series forecasting? [C] // Proceedings of the 38th Conference on Neural Information Processing Systems, 2024: 60162-60191.
- [24] Li WB, Yao D, Zhao RB, et al. STBench: assessing the ability of large language models in spatio-temporal analysis [Z/OL]. arXiv Preprint, arXiv: 2406.19065, 2024.
- [25] Gruver N, Finzi M, Qiu SK, et al. Large language models are zero-shot time series forecasters [C] // Proceedings of the 37th Conference on Neural Information Processing Systems, 2023: 19622-19635.
- [26] Bai YT, Geng XY, Mangalam K, et al. Sequential modeling enables scalable learning for large vision models [C] // Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 22861-22872.
- [27] OpenAI. GPT-4 technical report [Z/OL]. arXiv Preprint, arXiv: 2303.08774, 2023.
- [28] Touvron H, Martin L, Stone K, et al. Llama 2: open foundation and fine-tuned chat models [Z/OL]. arXiv Preprint, arXiv: 2307.09288, 2023.
- [29] Xue H, Salim FD. PromptCast: a new prompt-based learning paradigm for time series forecasting [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 36(11): 6851-6864.
- [30] Liu X, Hu JF, Li Y, et al. UniTime: a language-empowered unified model for cross-domain time series forecasting [C] // Proceedings of the ACM Web Conference, 2024: 4095-4106.
- [31] Wang CS, Qi Q, Wang JY, et al. ChatTime: a unified multimodal time series foundation model bridging numerical and textual data [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2025: 12694-12702.
- [32] Zhou T, Niu PS, Wang X, et al. One fits all: power general time series analysis by pretrained LM [C] // Proceedings of the 37th Conference on Neural Information Processing Systems, 2023: 43322-43355.
- [33] Sun CX, Li HY, Li YL, et al. TEST: text prototype aligned embedding to activate LLM's ability for time series [Z/OL]. arXiv Preprint, arXiv: 2308.08241, 2023.
- [34] Chang C, Wang WY, Peng WC, et al. LLM4TS: aligning pre-trained LLMs as data-efficient time-series forecasters [J]. ACM Transactions on Intelligent Systems and Technology, 2025, 16(3): 1-20.
- [35] Jin M, Wang SY, Ma LT, et al. Time-LLM: time series forecasting by reprogramming large language models [Z/OL]. arXiv Preprint, arXiv: 2310.01728, 2023.
- [36] Cao DF, Jia FR, Arik SO, et al. TEMPO: prompt-based generative pre-trained Transformer for time series forecasting [Z/OL]. arXiv Preprint, arXiv: 2310.04948, 2023.
- [37] Jia FR, Wang K, Zheng YX, et al. GPT4MTS: prompt-based large language model for multi-modal time-series forecasting [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2024: 23343-23351.
- [38] Liu Y, Qin G, Huang XD, et al. AutoTimes: autoregressive time series forecasters via large language models [C] // Proceedings of the 38th Conference on Neural Information Processing Systems, 2024: 122154-122184.

- [39] Lee G, Yu WC, Shin K, et al. TimeCAP: learning to contextualize, augment, and predict time series events with large language model agents [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2025: 18082-18090.
- [40] Beneduce C, Lepri B, Luca M. Large language models are zero-shot next location predictors [J]. *IEEE Access*, 2025, 13(1): 77456-77467.
- [41] Nguyen TCA, Tao TM, Lee C, et al. TM-LLM: token-motion embedding fusion for large language model-based vessel trajectory prediction [C] // Proceedings of the 2024 15th International Conference on Information and Communication Technology Convergence, 2024: 115-120.
- [42] Yan YB, Wen HM, Zhong SR, et al. UrbanCLIP: learning text-enhanced urban region profiling with contrastive language-image pretraining from the web [C] // Proceedings of the ACM on Web Conference 2024, 2024: 4006-4017.
- [43] Chen ZH, Zheng LN, Lu C, et al. ChatGPT informed graph neural network for stock movement prediction [Z/OL]. arXiv Preprint, arXiv: 2306.03763, 2023.
- [44] Xu HJ, Gao Y, Hui Z, et al. Language knowledge-assisted representation learning for skeleton-based action recognition [J/OL]. *IEEE Transactions on Multimedia*, 2025: 1-16. <https://doi.org/10.1109/TMM.2025.3543034>.
- [45] Garza A, Challu C, Mergenthaler-Canseco M. TimeGPT-1 [Z/OL]. arXiv Preprint, arXiv: 2310.03589, 2023.
- [46] Goswami M, Szafer K, Choudhry A, et al. MOMENT: a family of open time-series foundation models [Z/OL]. arXiv Preprint, arXiv: 2402.03885, 2024.
- [47] Liu Y, Zhang HR, Li CY, et al. Timer: generative pre-trained Transformers are large time series models [Z/OL]. arXiv Preprint, arXiv: 2402.02368, 2024.
- [48] Rasul K, Ashok A, Williams AR, et al. Lag-Llama: towards foundation models for probabilistic time series forecasting [Z/OL]. arXiv Preprint, arXiv: 2310.08278, 2023.
- [49] Gao SH, Koker T, Queen O, et al. UniTS: a unified multi-task time series model [C] // Proceedings of the 38th Conference on Neural Information Processing Systems, 2024: 140589-140631.
- [50] Wang YH, Qiu YY, Chen P, et al. ROSE: register assisted general time series forecasting with decomposed frequency learning [Z/OL]. arXiv Preprint, arXiv: 2405.17478, 2024.
- [51] Ansari AF, Stella L, Turkmen C, et al. Chronos: learning the language of time series [Z/OL]. arXiv Preprint, arXiv: 2403.07815, 2024.
- [52] Woo G, Liu CH, Kumar A, et al. Unified training of universal time series forecasting Transformers [Z/OL]. arXiv Preprint, arXiv: 2402.02592, 2024.
- [53] Li ZH, Xia L, Shi L, et al. OpenCity: open spatio-temporal foundation models for traffic prediction [Z/OL]. arXiv Preprint, arXiv: 2408.10269, 2024.
- [54] Yuan Y, Ding JT, Feng J, et al. A universal pre-training and prompting framework for general urban spatio-temporal prediction [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2025, 37(5): 2212-2225.
- [55] Man X, Zhang CH, Feng J, et al. W-MAE: pre-trained weather model with masked autoencoder for multi-variable weather forecasting [Z/OL]. arXiv Preprint, arXiv: 2304.08754, 2023.
- [56] Liu X, Liang YX, Huang C, et al. When do contrastive learning signals help spatio-temporal graph forecasting? [C] // Proceedings of the 30th International Conference on Advances in Geographic Information Systems, 2022: 1-12.
- [57] Li RF, Zhong T, Jiang XK, et al. Mining spatio-temporal relations via self-paced graph contrastive learning [C] // Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022: 936-944.
- [58] Lagler K, Schindelegger M, Böhm J, et al. GPT2: empirical slant delay model for radio space geodetic techniques [J]. *Geophysical Research Letters*, 2013, 40(6): 1069-1073.
- [59] Ma QL, Liu Z, Zheng ZJ, et al. A survey on time-

- series pre-trained models [J]. [IEEE Transactions on Knowledge and Data Engineering](#), 2024, 36(12): 7536-7555.
- [60] Li Z, Qiu X, Chen P, et al. TSFM-Bench: a comprehensive and unified benchmark of foundation models for time series forecasting [C] // Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2025: 5595-5606.
- [61] Goodge A, Ng WS, Hooi B, et al. Spatio-temporal foundation models: vision, challenges, and opportunities [Z/OL]. arXiv Preprint, arXiv: 2501.09045, 2025.
- [62] Ye JX, Zhang WQ, Yi K, et al. A survey of time series foundation models: generalizing time series representation with large language model [Z/OL]. arXiv Preprint, arXiv: 2405.02358, 2024.
- [63] Chang C, Wang WY, Peng WC, et al. Align and fine-tune: enhancing LLMs for time-series forecasting [Z/OL]. NeurIPS Workshop on Time Series in the Age of Large Models, 2024: 1-15.
- [64] Hu YX, Li Q, Zhang DX, et al. Context-alignment: activating and enhancing LLM capabilities in time series [Z/OL]. arXiv Preprint, arXiv: 2501.03747, 2025.